Int. J. Behav. Med. (2009) 16:89–90
DOI 10.1007/s12529-009-9049-1

EDITORIAL

Measurement Instruments in Behavioral Medicine

Joost Dekker

Published online: 11 June 2009

Measurements is the cornerstone of science. Measurement instruments (e.g., questionnaires, observational methods, and assessment of physiological parameters) are used to gather data: these data are used to test hypotheses (deductive approach) or to generate new hypotheses (inductive approach). In both approaches, a high-quality measurement instrument is an essential requirement for good science. This applies to behavioral medicine as well. It is therefore highly appropriate for the *International Journal of Behavioral Medicine* to publish the special series on 'measurement instruments in behavioral medicine'.

One of the challenges facing behavioral medicine is the bewildering variety of measurement instruments. Several years ago, 31 questionnaires and 26 performance-based tests were identified to measure pain and activity limitations in osteoarthritis [1, 2]. And this is osteoarthritis only: in other areas an equally wide range of instruments is available for measuring functioning and quality of life.

This wide range of measurement instruments constitutes a huge impediment to scientific communication. Results of studies using different instruments cannot be fully compared. Are questionnaire X and questionnaire Y really measuring the same construct or do they measure different concepts? How does a 15-point improvement on questionnaire X compare to a nine-point improvement on questionnaire Y? In a way, the situation in measuring health resembles the pre-Napoleonic era, when cities and even villages had their own measures of length, volume, and weight: this created confusion and impediments to science and trade [3]. It is high time for the field of behavioral medicine to reach consensus on preferred measurement instruments. In the natural sciences, consensus on the International System of Units was reached a little bit more than 150 years ago. Behavioral medicine should aim for a similar consensus.

At least one step into the direction of reaching consensus on preferred measurement instruments is the translation of existing questionnaires into other languages, instead of developing fully new questionnaires. Six papers in the present special series concern translation of an existing questionnaire. Ng et al. [4] evaluated a Chinese translation of the Daily Spiritual Experience Scale. Spindler et al. [5] and Pedersen et al. [6] translated the questionnaire for assessing Type D personality into the Danish and Ukrainian language, respectively. Canavarro et al. [7] translated the World Health Organization Quality of Life scale into the Portuguese language, while Maïano et al. [8] translated the Body Image Avoidance Questionnaire into the French language. In all cases, the psychometric properties of the translated questionnaires were satisfactory: translating existing instruments is an excellent alternative to developing new instruments.

However, translation might introduce subtle differences in various language versions of the same questionnaire. Even if the measurement properties of the various versions of the questionnaire are good, there is no guarantee for full measurement equivalence across cultures. In a carefully designed study, Choi et al. [9] demonstrated differences in item scores of the Job Content Questionnaire between European cultures. To reduce cross-language differences in the future, the authors argue in favor of a stricter translation process and studies using only items which have been shown not to suffer from cross-cultural differences.

J. Dekker (✉)
Department of Rehabilitation Medicine, EMGO Institute,
VU University Medical Centre,
P.O. Box 7057,
1007 MB Amsterdam, Netherlands
e-mail: j.dekker@vumc.nl

Three other papers in the special series concern the further validation of measurement instruments. Lee et al. [10] validated the Life Control Scale in a new population. Guo et al. [11] and Weck et al. [12] evaluated the factor structure of the Decisional Balance Scale and the Illness Attitude Scales, respectively. Cross-validation and evaluation of the factor structure of existing instruments are highly relevant objectives: in the context of reaching consensus on measurement instruments thorough information on measurement properties of measurement instruments is highly desirable.

The present special series is intended to contribute to information on properties of measurement instruments, in the context of reaching consensus on preferred measurements in behavioral medicine.

References

1. Veenhof C, Bijlsma JW, van den Ende CH, van Dijk GM, Pisters MF, Dekker J. Psychometric evaluation of osteoarthritis questionnaires: a systematic review of the literature. Arthritis Rheum. 2006;55(3):480–92.
2. Terwee CB, Mokkink LB, Steultjens MP, Dekker J. Performance-based methods for measuring the physical function of patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. Rheumatology (Oxford). 2006;45 (7):890–902.
3. Dekker J, Dallmeijer AJ, Lankhorst GJ. Clinimetrics in rehabilitation medicine: current issues in developing and applying measurement instruments 1. J Rehabil Med. 2005;37(4):193–201.
4. Ng SM, Fong TC, Tsui EY, Au-Yeung FS, Law SK. Validation of the chinese version of underwood's daily spiritual experience scale-transcending cultural boundaries? Int J Behav Med 2009. doi:10.1007/s12529-009-9045-5.
5. Spindler H, Kruse C, Zwisler AD, Pedersen SS. Increased anxiety and depression in Danish cardiac patients with a type d personality: cross-validation of the type D scale (DS14). Int J Behav Med 2009. doi:10.1007/s12529-009-9037-5.
6. Pedersen SS, Yagensky A, Smith OR, Yagenska O, Shpak V, Denollet J. Preliminary evidence for the cross-cultural utility of the type D personality construct in the Ukraine. Int J Behav Med 2009. doi:10.1007/s12529-008-9022-4.
7. Canavarro MC, Serra AV, Simoes MR, Rijo D, Pereira M, Gameiro S, et al. Development and psychometric properties of the world health organization quality of life assessment instrument (WHOQOL-100) in Portugal. Int J Behav Med 2009. doi:10.1007/s12529-008-9024-2.
8. Maiano C, Morin AJ, Monthuy-Blanc J, Garbarino JM. The body image avoidance questionnaire: assessment of its construct validity in a community sample of french adolescents. Int J Behav Med 2009. doi:10.1007/s12529-009-9035-7.
9. Choi BK, Bjorner JB, Ostergren PO, Clays E, Houtman I, Punnett L, et al. Cross-language differential item functioning of the job content questionnaire among European countries: the JACE Study. Int J Behav Med 2009. doi:10.1007/s12529-009-9048-2
10. Lee C, Ford J, Gramotnev H. The Life Control Scale: Validation with a Population Cohort of Middle-Aged Australian Women. Int J Behav Med 2009. doi:10.1007/s12529-008-9013-5.
11. Guo B, Aveyard P, Fielding A, Sutton S. The Factor Structure and Factorial Invariance for the Decisional Balance Scale for Adolescent Smoking. Int J Behav Med 2009. doi:10.1007/s12529-008-9021-5.
12. Weck F, Bleichhardt G, Hiller W. The Factor Structure of the Illness Attitude Scales in a German Population. Int J Behav Med 2009. doi:10.1007/s12529-009-9043-7.